Support Vector Regression (2)

Manel Martínez-Ramón

ECE, UNM

October 2018

- The ν-SVR is a way to automatically tune ε by adding a parameter that imposes a bound on the fraction of support vectors (Smola, 2003). The primal problem becomes

$$\text{Minimize } ||\mathbf{w}||^2 + C\left(\sum_{n=1}^{N}(\xi_i + \xi_i^*) + N\nu\varepsilon\right)$$

$$\text{subject to } \begin{cases} y_n - \mathbf{w}^\top \mathbf{x}_n - b - \varepsilon - \xi_n \leq 0 \\ -y_n + \mathbf{w}^\top \mathbf{x}_n + b - \varepsilon - \xi_n^* \leq 0 \\ \xi_n, \xi_n^* \geq 0 \end{cases}$$

- Here we introduce ε as part of the optimization, multiplied by a constant $N\nu$ which represents a fraction of the number of samples (where $0 < \nu < 1$ is another free parameter).

- The Lagrange optimization is equal to the one of the SVR, where the dual to be maximized is

$$L_d = -\frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^\top \mathbf{K}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*) + (\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)^\top \mathbf{y}$$

$$\text{subject to } \begin{cases} \sum_{n \in N_{sv}} (\alpha_n - \alpha_n^*) = 0 \\ \sum_{n \in N_{sv}} (\alpha_n + \alpha_n^*) = CN\nu \\ 0 \leq \alpha_n, \alpha_n^* \leq C \end{cases}$$

where N_{sv} is the number of support vectors.

- Dual constraint

$$\begin{cases} \sum_{n \in N_{sv}} (\alpha_n + \alpha_n^*) = CN\nu \\ 0 \leq \alpha_n, \alpha_n^* \leq C \end{cases}$$

mean that ν **is a lower bound on the fraction of support vectors**.

- Indeed, assume that all support vectors are outside the ϵ-tube. Then, $\alpha_n, \alpha_n^* = C$ for all of them. Then

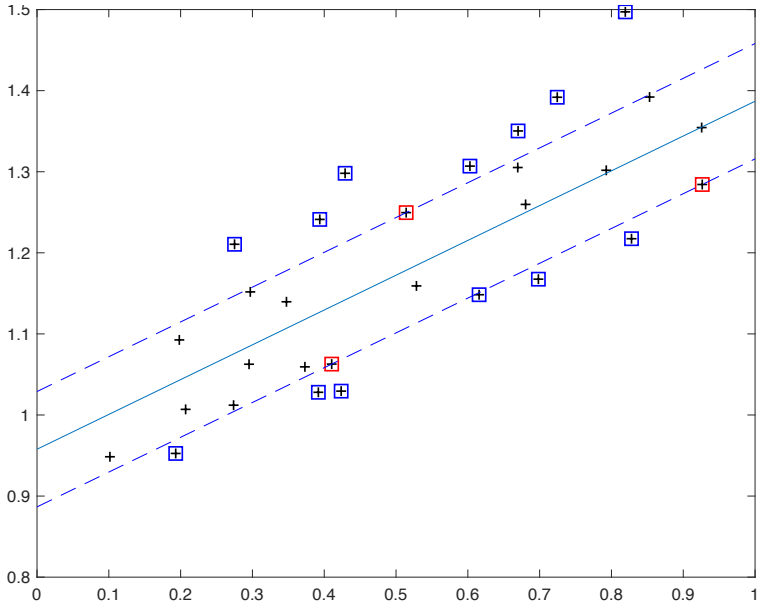
$$\sum_{n \in N_{sv}} (\alpha_n + \alpha_n^*) = CN_{sv}$$

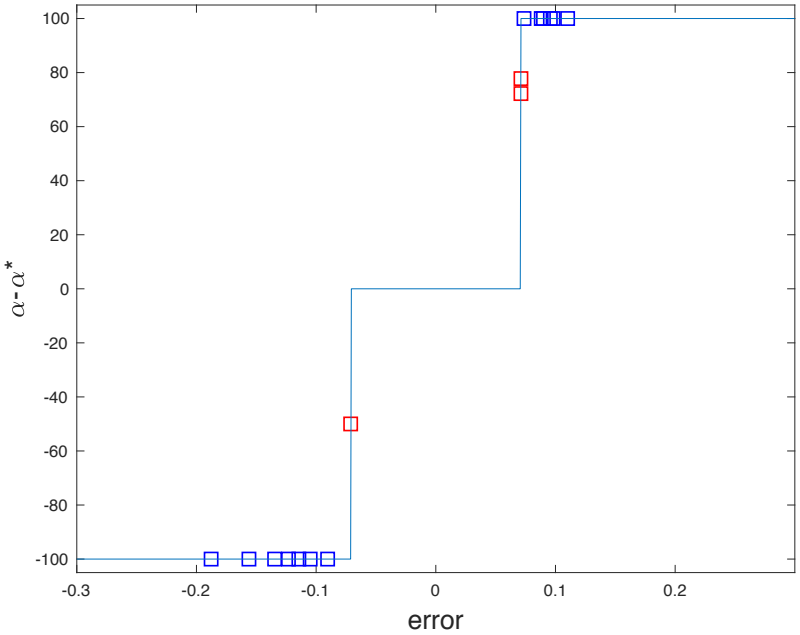
Thus $N_{sv} = \nu N$.

- If there are support vectors on the margin, the number of support vectors will be higher.

$$N_{sv} \geq \nu N$$

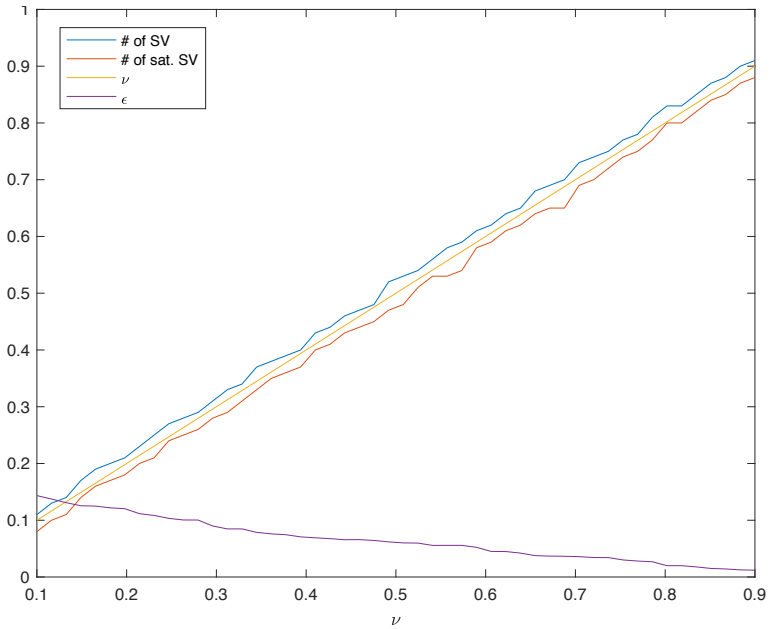
- In either case the number of support vectors outside the margin will be equal or less than νN: Hence, ν **is an upper bound of the number of errors of absolute value higher than** ϵ.





One dimension example of regression using ν-SVR with $\nu = 0.5$. The total number of SV is 16, and the number of saturated ones is 13, of a total of 30 training data.

Examples



In this example, the value of ν is swept from 0.1 to 0.9. The margin (decreasing line) is higher when the fraction of support vectors is lower. The straight line represents ν. The blue (upper line) represents the fraction of total SVs, which is always higher than ν. The red (lower) line represents the fraction of non saturated support vectors, which is always less than ν.

Outcomes of this lesson

We have seen here the ν-SVM (nu-SVM). After it you must have a clear idea of:

- The purpose of the modification that introduces ν.
- The meaning of the constraints found in the dual.
- The behaviour of the machine with respect to ν.